Investigating the Relationship Between Deprivation and Criminal Damage in Westminster: Insights from Regression Analysis

Research Question

The purpose of this research is to examine the relationship between deprivation and criminal damage by analysing data on deprivation and recorded criminal damage in Westminster, spanning from December 2020 to November 2023. The primary question being addressed is: Is there a correlation between deprivation and criminal damage? Additionally, this study seeks to explore whether underlying indicators of deprivation have an impact on criminal damage and to quantify this effect. This will provide a detailed understanding of the relationship between Deprivation and Criminal Damage in Westminster.

Literature Review

An important topic related to every society - deprivation has been studies for decades. It is a multifaceted concept, a state where a person is noticeably at a lesser advantage compared to others in their community or society, and can be material or social (Townsend, 1987). Moreover, this definition also extends to psychological and health-related dimensions (Fu, Exeter and Anderson, 2015).

The definition of criminal damage in the UK involves the intentional or reckless destruction or damage of property belonging to another (Criminal Damage Act 1971, no date). This type of crime - criminal damage can create a perception of community instability and neighborhood deterioration, influencing both the local economy and community well-being (Gibbons, 2004). Meanwhile, Westminster has the highest crime rate in London, with Criminal Damage rates 45% above average (Westminster Crime and Safety Statistics, no date). Also, it ranks in the top 30 percent of deprived local authorities nationally (English indices of deprivation 2019, no date). De Courson and Nettle (2021) suggest that deprivation leads to high crime rates. Thus, quantifying deprivation and probing its link with Criminal damage is crucial research for enhancing social welfare and understanding how to introduce policies that indirectly mitigate the community impact of Criminal damage by reducing deprivation.

In this case, the effect of deprivation on criminal behavior, particularly in urban areas, has been noted. Many studies have aimed to quantify and interpret their relationship over the years. A study by Sariaslan *et al.*, (2013) indicated that deprivation had a significant but complex influence on criminal behavior by employing generalized linear mixed-effects models. Ciacci and Tagliafico (2020) employed two aggregate quantitative analyses to build indices, where crime is measured through a summation of standardized variables, and social deprivation is assessed using the DP2 method. Meanwhile, quantitative analysis exploring the correlation between deprivation and the specific crime type of criminal damage appear unattempted.

However, quantifying deprivation is a challenging task. Townsend (1987) , Carstairs and Morris (1990) each proposed a method for calculating deprivation score, but their approaches focused solely on material disadvantage. Moreover, initially developed to assess workloads in general practice (Jarman, 1991), the Jarman score eventually evolved into an indirect indicator of deprivation, but it has not been recalculated since the Census 2001. Additionally, the UK government first released Indices of Deprivation (IoD) starting in 2000, which started to conceptualize multiple deprivations at a small-area level as an accumulation of discrete dimensions or domains of deprivation (English indices of deprivation 2019, no date). These indices have been fundamental in targeting policy in the UK (Noble et al., 2006).

Data

Two datasets were utilised in this investigation. Crime data was sourced from data.police.uk, the UK's official crime and policing open data site. Spanning December 2020 to November 2023, the anonymised data includes over a million records from the Metropolitan Police Service. Each crime record includes information shown in Table 1. The filtered criminal damage count distribution of Westminster is shown in Fig. 1.

Deprivation data came from the most recent English Indices of Deprivation (IoD) 2019, a set of relative measures of deprivation for small areas or neighbourhoods, published by the Ministry of Housing, Communities & Local Government. The IoD 2019 utilizes 39 indicators across seven domains. These indicators are combined, using a specific weighting system, with domains weighted as shown in Fig. 2 to produce the Index of Multiple Deprivation (IMD) 2019 – a measure of deprivation affecting individuals in each area, calculated for each Lower-layer Super Output Area (LSOA). Note that it measures relative deprivation, therefore cannot measure the absolute scale.

Our research uses the underlying indicators from the dataset, which contribute to the construction of the IMD, as potential explanatory variables for deprivation. The data sources for each indicator are as up-to-date as possible, though there's no single consistent time point for all of them due to varying availability. This combined with the fact that the offenders may not reside locally, could introduce variability and inconsistencies in our regression model and might also cause multicollinearity due to the many indicators.

Table 1: Crime record fields and meanings.

Field	Meaning
Reported by	The crime reported by force.
Falls within	The crime falls within the Lower Layer Super Output Area.
Longitude and **Latitude**	The anonymised coordinates of the crime.
LSOA code and **LSOA name**	References to the Lower Layer Super Output Area that the anonymised point falls into.
Crime type	One of the crime types listed in the Police.
Last outcome category	A reference to whichever of the outcomes associated with the crime occurred most recently.
Context	A field provided for forces to provide additional human-readable data about individual crimes. (always empty)

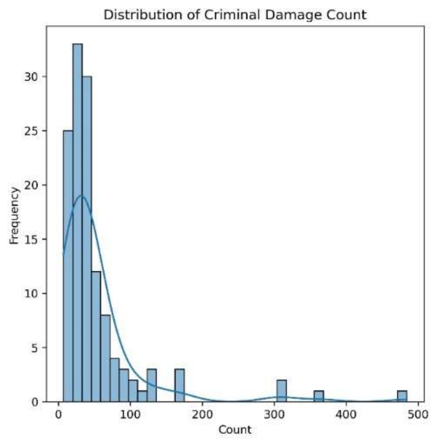

Figure 1: Distribution of Criminal Damage Count in Westminster from December 2020 to November 2023

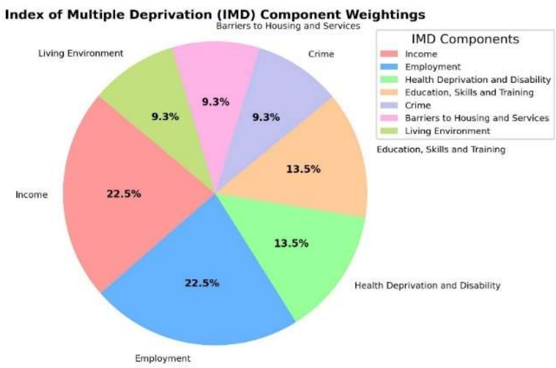

Figure 2: Index of Multiple Deprivation (IMD) Component Weightings

Methodology

We employed Spearman's rank correlation analysis for IMD and Criminal Damage Count, considering that IMD data are ordinal thus fit Spearman's applicability. We set the Significance level (α) = 0.05. Then the p-value = 0.472. With the p > α, indicates the result is NOT SIGNIFICANT. Therefore, there is no significant correlation between IMD and criminal damage in Westminster.

However, to explore the relationship between IoD's 39 variables and Criminal Damage Count, we conducted a correlation analysis and obtained a correlation matrix as shown in Fig. 3.

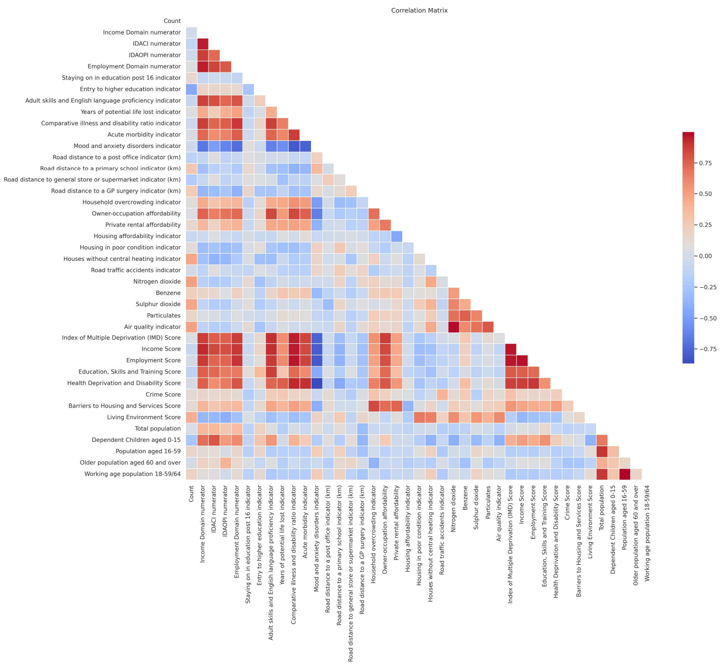

Figure 3: Correlation Matrix between Underlying Indicators and Criminal Damage Count

The results indicate that many variables within the dataset display correlation with the Criminal Damage Count. Consequently, we attempt to construct a multiple linear regression model to attempt to find a possible explanatory model for the relationship between criminal damage and the underlying indicators. The model is described as follows:

$$Y = \beta_0 + \beta_{IncomeDomainNumerator} X_{IncomeDomainNumerator} + \cdots \\ + \beta_{Outdoors\ Sub-domain\ Score} X_{Outdoors\ Sub-domain\ Score} + \varepsilon$$

In this model, Y is the dependent variable, $X_{IncomeDomainNumerator}$ $\cdots$ $X_{Outdoors\ Sub-domain\ Score}$ represents all the independent variables, and β_0 , $\beta_{IncomeDomainNumerator}$ $\cdots$ $\beta_{Outdoors\ Sub-domain\ Score}$ are the parameters of the model, representing the intercept and the coefficients for each independent variable. ε is the error term representing random errors that the model does not account for. We employ the method of ordinary least squares (OLS) fitting to estimate the parameters, to minimize the sum of squared residuals between the predicted values from the model and the actual observed values.

However, to enhance the effectiveness of the model, we need to employ the Box-Cox transformation on both the dependent and independent variables. This is a statistical technique used for improving the normality, homoscedasticity, and linear relationship of the data (Box and Cox, 1964).

The one-parameter Box–Cox transformations are defined as:

$$y_i^{(\lambda)} = \begin{cases} \dfrac{y_i^{\lambda} - 1}{\lambda}, & \text{if } \lambda \neq 0; \\ \ln y_i\ , & \text{if } \lambda = 0. \end{cases}$$

Where y_i represents the i^{th} data point in the dataset. λ is the Box-Cox transformation parameter. $y_i^{(\lambda)}$ represents the transformed value of the i^{th} data point after applying the Box-Cox transformation.

Considering that the Box-Cox transformation requires the values to be positive (Box and Cox, 1964), we apply a shift to variables whose minimum value is less than or equal to zero. Following this shift, we then perform the Box-Cox transformation on all variables and the count of criminal damage. Figs. 4 and 5 shows the output before and after box-cox transformation.

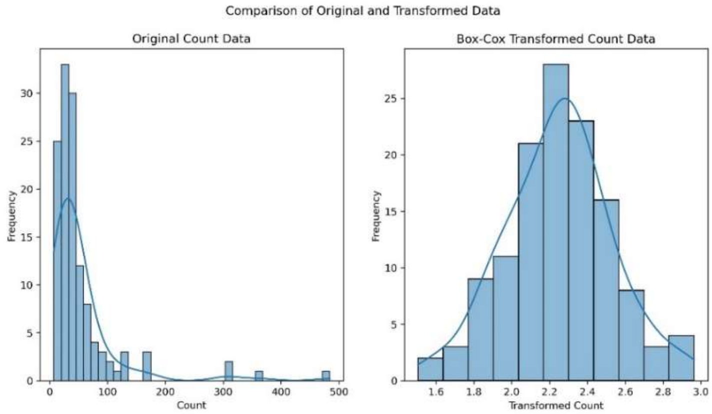

Figure 4: Comparison of Original and Transformed Count of Criminal Damage.

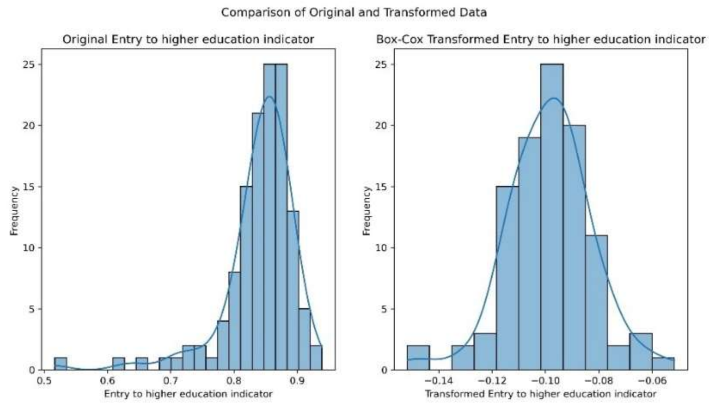

Figure 5: Comparison of Original and Transformed Higher Education Indicator as an example of the 39 variables.

According to Fig. 3, multicollinearity is a significant issue. It refers to a situation in a regression model where there is a high correlation or linear dependency among the independent variables, leading to instability in parameter estimation and difficulties in interpretation. We can utilize the Variance Inflation Factor (VIF) to measure the degree of multicollinearity between each of the independent variables (University College London, 2023).

Given x_1, x_2, …, x_p, the VIF for the x_k variable is:

$$VIF_k = \frac{1}{1 - R_k^2}$$

where R_k^2 is the R^2 value obtained by regressing the x_k on the remaining x variables: $x_k \sim x_1 + \cdots + x_{k-1} + x_{k+1} + \cdots + x_p + b$, for k=1, 2, …, p.

A larger VIF_k, the higher multicollinearity. In this case, we will filter out the variables with a VIF > 5. However, interpreting the model becomes complex after the Box-Cox transformation. Therefore, based on the definition of the Box-Cox transformation and our linear regression model, for the cases where $\lambda \neq 0$, we can calculate that when $X'_{i,0}$ changes by t:

$$\Delta Y = \left(\beta_0 + \beta_1 X'_1 + \cdots + \beta_i\left(X'_{i,0} + t\right) + \cdots + \beta_n X'_n)\lambda_Y + 1\right)^{\frac{1}{\lambda_Y}}$$

$$- \left(\beta_0 + \beta_1 X'_1 + \cdots + \beta_i X'_{i,0} + \cdots + \beta_n X'_n)\lambda_Y + 1\right)^{\frac{1}{\lambda_Y}}$$

Where ΔY is the change in the dependent variable Y on the original scale. β_0 is the intercept, $\beta_1, \beta_2, \dots, \beta_n$ are the coefficients of each variable. $X'_1, X'_2, \dots, X'_n$ represent the values of the independent variables in the model. $X'_{i,0}$ is the initial value of the i^{th} independent variable X_i before it changes. t is the change in the independent variable X_i .

Residuals are the differences between the observed values and the values predicted by our model. Ideally, the residuals should be randomly scattered around 0, with no discernible pattern in the Residual vs. Fitted Values Plot and the points should fall approximately along a straight line in the Normal Q-Q Plot, meaning that the residuals are approximately normally distributed.

Results

Table 2: The output of the explanatory variables

	Coefficient	Std. Error	T	P > |t|
Staying on in education post 16 indicator	-0.3653	0.135	-2.706	0.008
Barriers to Housing and Services Score	-0.0276	0.012	-2.261	0.026
Houses without central heating indicator	0.3337	0.083	4.023	0.000
constant	6.8701	4.102	1.675	0.038

Table 3: The output of the explanatory model

R^2	Adj. R^2	F	Prob >F
0.422	0.353	2.505	0.00139

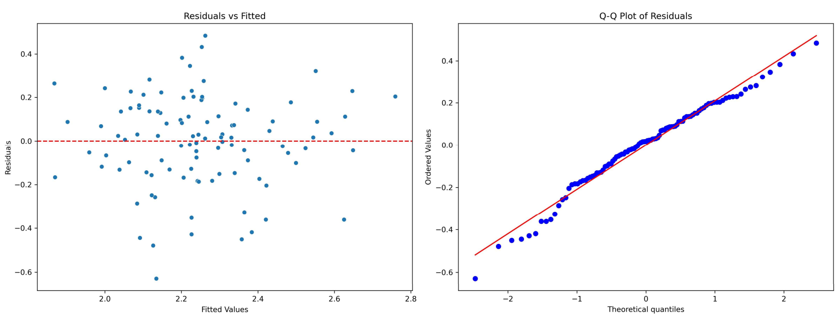

Figure 6: Residual vs. Fitted Values Plot and Q-Q Plot of Residuals

Discussion

The regression result shows that 35.3% of the variance in the transformed criminal damage count can be explained (Adjusted R^2 in Table 3). Fig. 6 shows a good model fit, evidenced by the residual patterns. According to Table 2, the p-values signify that three explanatory variables are significant: the 'Staying on in education post-16 indicator,' the 'Houses without central heating indicator,' and the 'Barriers to Housing and Services Score.' When each of these increases by 1, based on the aforementioned formula for calculating the ΔY, the original Y, which is the Criminal Damage Count, will increase by 0.292, decrease by 0.161, and increase by 0.017, respectively.

According to the IoD 2019 interpretations of the variables and the results, as the proportion of adolescents staying in education beyond the age of 16 increases, incidents of Criminal Damage in the area also rise, potentially due to more youths congregating, seeking belonging or avenues for self-expression (Hart, 2009). Additionally, an increase in Barriers to Housing and Services may lead to community discontent and tension, potentially driving individuals to unlawful means for sustaining livelihoods (McGahey, 1986). Conversely, a correlation exists between an increase in houses without central heating and a decrease in Criminal Damage, possibly due to the migration of younger populations, who more likely to commit criminal damage, in search of better living conditions (Sweeten, Piquero and Steinberg, 2013).

Conclusion

Our analysis shows no correlation between IMD and Criminal Damage, but a link exists between deprivation, represented by three IoD 2019 variables, and Criminal Damage. Although the causality behind this relationship worth further investigation, a possible explanatory model to simulate the count of Criminal Damage in Westminster has been produced. However, with an Adjusted R^2 of only 35.3%, the explanation power of the model is relatively limited, hinting at other complex factors influencing crime beyond deprivation dimension. Furthermore, the deprivation quantification method could improve by incorporating additional factors. From a policy-maker’s perspective, upgrading the education system, reinforcing basic law education over juveniles, and providing housing and basic living subsidies could potentially be beneficial for mitigating occurrences of Criminal Damage.

Word Count (excluding References):	*1746*

References

Box, G. E. P. and Cox, D. R. (1964). ‘An Analysis of Transformations’. *Journal of the Royal Statistical Society: Series B (Methodological)*, 26 (2), pp. 211–243. doi: 10.1111/j.2517-6161.1964.tb00553.x.

Carstairs, V. and Morris, R. (1990). ‘Deprivation and health in Scotland’. *Health bulletin*, 48 (4), pp. 162–175.

Ciacci, A. and Tagliafico, G. (2020). ‘Measuring the Existence of a Link between Crime and Social Deprivation within a Metropolitan Area’. Universidad Sevilla. doi: 10.12795/rea.2020.i40.04.

De Courson, B. and Nettle, D. (2021). ‘Why do inequality and deprivation produce high crime and low trust?’ *Scientific Reports*. Nature Publishing Group, 11 (1), p. 1937. doi: 10.1038/s41598-020-80897-8.

English indices of deprivation 2019. (no date). *GOV.UK*. Available at: https://www.gov.uk/government/statistics/english-indices-of-deprivation-2019 (Accessed: 15 January 2024).

Fu, M., Exeter, D. and Anderson, A. (2015). ‘The politics of relative deprivation: A transdisciplinary social justice perspective.’ *Social science & medicine*, 133, pp. 223–32. doi: 10.1016/j.socscimed.2014.12.024.

Hart, S. (2009). ‘The “problem” with youth: young people, citizenship and the community’. *Citizenship Studies*. Routledge, 13 (6), pp. 641–657. doi: 10.1080/13621020903309656.

Jarman, B. (1991). ‘Jarman index.’ *BMJ : British Medical Journal*, 302 (6782), pp. 961–962.

McGahey, R. M. (1986). ‘Economic Conditions, Neighborhood Organization, and Urban Crime’. *Crime and Justice*. The University of Chicago Press, 8, pp. 231–270. doi: 10.1086/449124.

Participation, E. (no date). *Criminal Damage Act 1971*. Statute Law Database. Available at: https://www.legislation.gov.uk/ukpga/1971/48/contents (Accessed: 15 January 2024).

Sariaslan, A., Långström, N., D’Onofrio, B., Hallqvist, J., Franck, J. and Lichtenstein, P. (2013). ‘The impact of neighbourhood deprivation on adolescent violent criminality and substance misuse: A longitudinal, quasi-experimental study of the total Swedish population’. *International Journal of Epidemiology*, 42 (4), pp. 1057–1066. doi: 10.1093/ije/dyt066.

Sweeten, G., Piquero, A. R. and Steinberg, L. (2013). ‘Age and the Explanation of Crime, Revisited’. *Journal of Youth and Adolescence*, 42 (6), pp. 921–938. doi: 10.1007/s10964-013-9926-4.

Townsend, P. (1987). ‘Deprivation’. *Journal of Social Policy*. Cambridge University Press, 16 (2), pp. 125–146. doi: 10.1017/S0047279400020341.

Westminster Crime and Safety Statistics | CrimeRate. (no date). Available at: https://crimerate.co.uk/london/westminster (Accessed: 16 January 2024).